INDEXING OF MID-RESOLUTION SATELLITE IMAGESWITH STRUCTURAL ATTRIBUTES

A. Bhattacharya[a,b], M. Roux[a], H. Maître[a], I. Jermyn [b], X. Descombes [b], J. Zerubia [b]

[a]Institut TELECOM, TELECOM ParisTech, CNRS UMR 5141 LTCI,75013 Paris, France - (avik.bhattacharya, michel.roux, henri.maitre)@telecom-paristech.fr
[b]Ariana (joint research group INRIA/I3S) INRIA, BP 93, 06902 Sophia Antipolis, Cedex France - (Ian.Jermyn, Xavier.Descombes, Josiane.Zerubia)@inria.fr

Commission IV, WG IV/2

KEY WORDS: Landscape, Segmentation, Features, Extraction, Classification, Modelling, Data Mining

ABSTRACT:

Satellite image classification has been a major research field for many years with its varied applications in the field of Geography, Geology, Archaeology, Environmental Sciences and Military purposes. Many different techniques have been proposed to classify satellite images with color, shape and texture features. Complex indices like Vegetation index (NDVI), Brightness index (BI) or Urban index (ISU) are used for multi-spectral or hyper-spectral satellite images. In this paper we will show the efficiency of structural features describing man-made objects in mid-resolution satellite images to describe image content. We will then show the state-of-the-art to classify large satellite images with structural features computed from road networks and urban regions extracted on small image patches cut in the large image. Fisher Linear Discriminant (FLD) analysis is used for feature selection and a one-vs-rest probabilistic Gaussian kernel Support Vector Machines (SVM) classification method is used to classify the images. The classification probabilities associated with each subimage of the large image provide an estimate of the geographical class coverage.

1. INTRODUCTION

The growth of large image databases during the last few decades with the advancement in image acquisition technologies have attracted researchers from different fields to work in the domain of image information mining systems. These images coming from various sources must be systematically analyzed to render important information which are often less relevant to human perception. The technologically advanced satellite sensors and the new storage systems have made image data too vast and complex. The manual annotation to describe a complex image completely is not feasible. Indexing and retrieval from remote sensing image databases relies on the extraction of appropriate information from the data about the entity of interest (Daschiel and Datcu, 2005). Indexing satellite images (Maitre, 2007) depends on the choice of features which in turn are dependent on the type and resolution of the sensors. For instance SIFT descriptors are widely used in the domain of multimedia (Lowe, 2004). Complex indices like Vegetation index (NDVI), Brightness index (BI) or Urban index (ISU) are used for multi-spectral or hyper-spectral images. Texture features are known to be highly discriminative for low resolution panchromatic images (Schroeder et al., 1998). Structural features describing man-made objects in mid-resolution images are most efficient to describe image content (Bhattacharya et al., 2007). The road network contained in an image is one example. The properties of road networks vary considerably from one geographical environment to another. The structural features computed from them can therefore be used to classify and retrieve such environments (Bhattacharya et al., 2007). In order to compute the structural features of the road network, we first need to extract the road network from the image and then convert the output to an appropriate representation. This representation must be absolutely independent of any extraction method. The road extraction methods are in general resolution dependent. An optimal road network extraction algorithm to accurately delineate road structures for all practical purposes is very hard to achieve. The methods used in our study are robust on many such road characteristics but they often failed to extract the narrow and finely structured road networks which are almost hidden in small urban areas. This failure of the extraction methods and hence the features computed from road networks poorly classify images containing such areas. In order to obtain some meaningful information from these regions, we need to segment such areas occurring in the images. A new set of structural features computed on segmented urban areas combined with the existing road network features provided an improved classification of the geographical environments.

In images, pixels provide the most basic level of information. The pixel values are the measurements of the satellite sensors of a region on the Earth surface. The information from these pixels are at a level far below the semantic meaning of the desired object or region. The classification of images based on the pixel values is tedious and expensive and hence is not an efficient strategy. In this paper we present a novel methodology to classify large satellite images with patches of images extracted from them. This is a novel idea in the sense that the patches considered contain a significant coverage of a particular type of geographical class. A one-vs-rest probabilistic Gaussian kernel Support Vector Machines (SVM) classification method is used to classify the images. In the work presented in this paper we have defined 7 such classes. These classes can be categorized as follows: 2 urban classes consisting of “Urban USA” and “Urban Europe”; 3 rural classes consisting of “Villages”, “Mountains” and Fields; an “Airports” class and a “Common” class (this can be considered as a rejection class indicating in particular images from seas).

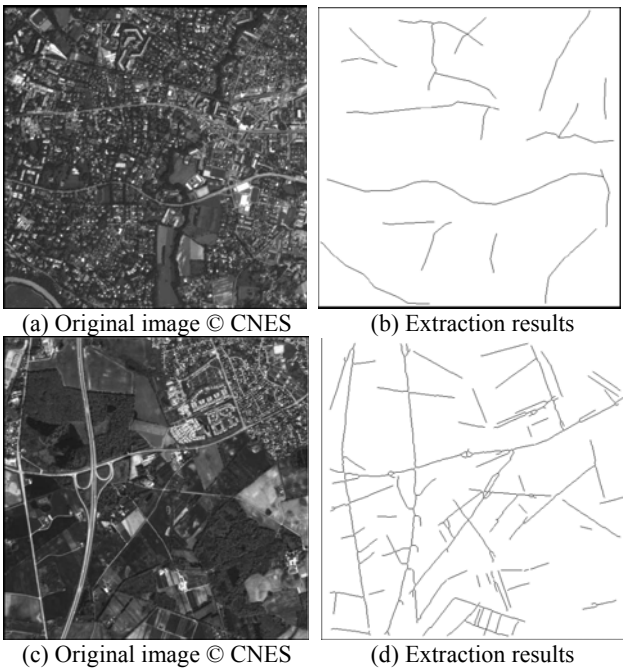
2. STRUCTURAL ATTRIBUTES

In this section we present four road network extraction methods, two of which were used in this work. A method is proposed to represent the extracted road networks as graphs. A morphological segmentation method is proposed to segment the urban areas in an image. Finally we describe the road network and urban region structural features used for the classification and indexing of satellite images.

2.1 Road Network and Urban Region Extraction

In order to compute the structural features of the road network, we first need to extract the road network from the image, and then convert the output to an appropriate representation. This representation should be independent of the output of the extraction algorithm, since we do not want to be committed to any single such method. In the preliminary studies reported in (Bhattacharya et al., 2006) we considered two network extraction methods (Rochery et al., 2003, Lacoste et al., 2005). The method of (Rochery et al., 2003) is based on Higher-Order Active Contours (HOACs) which are a generalization of standard active contours. The method of (Lacoste et al., 2005) models the line network as an object process, where the objects are interacting line segments. The output is a set of line segments of varying lengths, orientations, and positions. In spite of producing good extraction results, these methods were not used in this work due to the fact that they were not adapted for an optimization on large image databases, since manual expertise is needed to set the parameters in the algorithms according to image complexities.

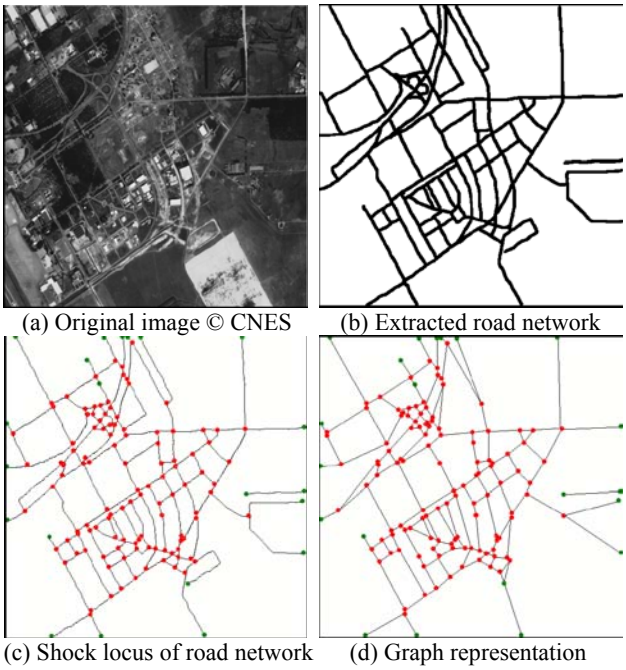
In the work reported in this paper we considered the two network extraction methods reported in (Fischler et al., 1981, Desolneux et al., 2000). These methods were rather easier to handle and could easily be adapted to large image database. The parameters once set in the algorithms works well with images of certain resolution. The output of the method described in (Fischler et al., 1981) is a binary image, which after a distance function computation can serve directly as an input to our method. Figure 1(b) shows examples of the extracted network. The output of the method described in (Desolneux et al., 2000) is a list of multiply aligned segments. In order to have a suitable input for our method, we convert the output of this method into a binary image, and use some image processing techniques to obtain single connected segments. Figure 1(d) shows examples of the extracted network. We then compute a distance function which then acts as an input to our method. The distance function resulting from these methods is converted to a graph representation of the road network for feature computation purposes. The graph itself captures the network topology, while the network geometry is encoded by decorating the vertices and edges with geometrical information. The conversion is performed by computing the shock locus of the distance function using the method of (Dimitrov et al., 2000, Siddiqi et al., 2002), extended to deal with multiple and multiply connected components with the depth-first search (DFS) algorithm (Cormen et al., 2001). The method identifies the shock points by finding out the limiting behavior of the average outward flux of the distance function as the region enclosing the shock point shrinks to zero. A suitable thresholding on this flux yields an approximation to the shock locus. The graph is constructed by taking triple (or, exceptionally, higher degree) points and end points as vertices, corresponding to junctions and terminals, while the edges are composed of all other points, and correspond to road segments between junctions and terminals. Figure 2 shows an example of the representation graph. The road network, Figure 2(b) is first extracted from the input image Figure 2(a).



(a) Original image © CNES (b) Extraction results

(c) Original image © CNES (d) Extraction results

Figure 1: Extraction results with 2 methods. Example (b) is with the method of (Fischler et al., 1981) and example (d) is with themethod of (Desolneux et al., 2000)



(a) Original image © CNES (b) Extracted road network

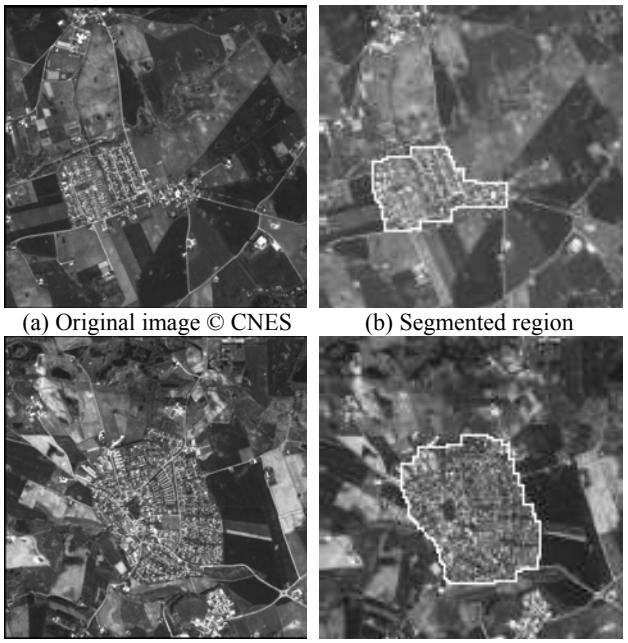
(c) Shock locus of road network (d) Graph representation

Figure 2: An example of the graph representation.

The methods cited in our study are robust on many such road characteristics but they often failed to extract the narrow and finely structured road networks which are almost hidden in small urban areas. This failure of the extraction methods and hence the features computed from road networks poorly classify images containing such areas. In order to obtain some

meaningful information from these regions, we need to segment such areas occurring in the images.

The heterogeneity and the geometrical complexity of urban structures in low radiometric and mid-resolution (2m or 5m) images show textural effects for objects with few pixel width. In our study we use the work of (Roux, 1992) developed to extract the urban regions from SPOT images. In SPOT images, the urban zones appear to be strongly textured and the problem of extraction of the regions is essentially a problem of differentiation of textures. The method used here is inspired from the works of (Serendero, 1989) and (Khatir, 1989). The principle idea is to extract the zone of high density of light and dark peaks. The techniques used are of mathematical morphology operations of opening and closing. The segmented compact urban regions in the images are shown in Figure 3(b) and Figure 3(d).



(a) Original image © CNES (b) Segmented region

(c) Original image © CNES (d) Segmented region

Figure 3: Images containing small urban areas and their segmentations.

2.2 Road Network Features

In this section we focus on 16 features summarized in table 1. These features can be categorized into six groups: six measures of 'density', four measures of 'curviness', two measures of 'homogeneity', one measure of 'length', two measures of 'distribution' and one measure of 'entropy'. We will now define the road network features.

Let v be a vertex and e be an edge. Let l_e be the length of the road segment corresponding to e, and let d_e be the length of e, that is the Euclidean distance between its two vertices. Let m_v be the number of edges at a vertex. Then $N_J = \sum 1_{v, m_v > 2}$ is the number of junction vertices and $E_J = \sum_{m_v > 2} m_v$ is the number of junction edges. Let Ω be the area of the image in pixels. We define the 'junction density' to be $\tilde{N}_J = \Omega^{-1} N_J$ and 'density of junction edges' to be $\tilde{E}_J = \Omega^{-1} E_J$. These are intuitively a useful measure to separate urban and rural areas: we expect urban areas to have a higher value of $\tilde{N}_J$ and $\tilde{E}_J$ than rural areas. Similarly, we define the 'network length' $L = \sum_e l_e$ and the 'length density' to be $\tilde{L} = \Omega^{-1} L$. Again, we expect urban areas to have a higher value of $\tilde{L}$ than rural areas. Note than one can have a high value of $\tilde{L}$ and a low value of $\tilde{N}_J$ if junctions are complex and the road segments are 'space-filling'. We also compute the network area Ω_L as the number of pixels corresponding to the network from the extracted binary image and define the 'network area density' as $\tilde{A} = \Omega^{-1} \Omega_L$. As can be seen in figure 2, many junction points are clustered around a small area in the network. To obtain a local characteristic of the junction density, we define a measure called 'local junction': $\tilde{N}_{r,j} = \Omega_{j,r}^{-1} \sum 1_{v \in \Omega_{j,r}, m_v > 2}$. This is the density of junction points falling in a circular region of radius r centered at junction point j. We then compute the mean and the variance of these junction densities over all junction points, mean($\tilde{N}_{r,j}$) and var($\tilde{N}_{r,j}$). A high var($\tilde{N}_{r,j}$), indicates the sparse structure of road junctions. Rural network structures will show such a characteristic. A low value indicates that junction points are clustered close to many other junction points, which is a prominent measure of urban network structure. The mean($\tilde{N}_{r,j}$) is also used as a measure of density.

Let $p_e = l_e / d_e$, and $k_e = l_e^{-1} \int_e |curv(s)| \, ds$, i.e., the absolute curvature per unit length of the road segment corresponding to the edge e. Although it may seem natural to characterize the network using the average values per edge of these quantities, in practice we have found that the variances of these quantities are equally useful. We thus define the 'ratio of lengths variance' and the 'ratio of lengths mean' to be the variance and mean of p_e over edges, var(p) and mean(p), and the 'average curvature variance' and 'average curvature mean' to be the variance and mean of k_e over edges, var(k) and mean(k). Note that it is quite possible to have a large value of p_e for an edge while having a small value of k_e if the road segment is composed of long straight segments, and vice-versa, if the road 'wiggles' rapidly around the straight line joining the two vertices in the edge. We expect rural areas to have high values of one of these two quantities, while urban areas will probably have low values, although this is less obvious than for the density measures.

To measure network homogeneity, we divide each image into four quadrants, labelled a. Subscript a indicates quantities evaluated for quadrant a rather than the whole image. Let $M_{J,a} = \sum_{v \in a, m > 2} m_v$ be the number of edges emanating from junctions in quadrant a. This is very nearly twice the number of edges in a, but it is convenient to restrict ourselves to junctions to avoid spurious termini at the boundary of the image. Let $\tilde{M}_{J,a} = \Omega^{-1} M_{J,a}$ be the density of such edges in quadrant a. Then we define the 'network inhomogeneity' to be the variance of $\tilde{M}_{J,a}$ over quadrants, var($\tilde{M}_J$). We also include mean($\tilde{M}_J$) as a feature.

In order to distinguish between the two urban classes (USA and Europe), the entropy of the histogram of angles at junctions, H_β, where β_j is the vector of angles between road segments at junction j, is a good measure. As is evident from the physical

characteristics of these road network structures, roads in USA tend to be parallel and cross each other orthogonally forming T-junctions or crossroads, whereas European roads tend to wiggle and meet or cross each other at roundabouts. Thus it seems natural that $H_\beta \leq 2$ bits are necessary to encode information about road segments at junctions for road networks in the USA, whereas for road networks in Europe, $H_\beta \geq 2$ bits are necessary. The same measure can also be used to distinguish between Mountains and Fields, while the 'density' features distinguish rural networks from urban networks.

A 'distribution' measure of edges at a vertex provides us with information as to how the edges at a vertex are distributed in the network. Let $E_{D,i}$ be the proportion of junction points with i edges at them. We use mean($E_{D,i}$) and var($E_{D,i}$) as features. The variance of the edge distribution is lower in the case of networks in urban areas as opposed to rural, and it is lower also in the case of urban networks in the USA as opposed to in Europe.

Notation	Description
$\tilde{N}_J$	**Junction density**
L	**Network length**
$\tilde{L}$	**Length density**
$\tilde{A}$	**Network area density**
p_e	Ratio of length
var(p)	**Ratio of lengths variance**
mean(p)	**Ratio of lengths mean**
k_e	Average curvature
var(k)	**Average curvature variance**
mean(k)	**Average curvature mean**
$E_{D,i}$	Number of junction with $m_v=i$
$var(E_{D,i})$	**Edge distribution variance**
$mean(E_{D,i})$	**Edge distribution mean**
E_J	Number of junction edges
$\tilde{E}_J$	**Junction edges density**
$\tilde{M}_{J,a}$	Density of junction edges per quadrant
$var(\tilde{M}_J)$	**Junction edges density variance**
$mean(\tilde{M}_J)$	**Junction edges density mean**
$\tilde{N}_{r,i}$	Local junction density
$var(\tilde{N}_{r,i})$	**Variance of the local junction densities**
$mean(\tilde{N}_{r,i})$	**Mean of the local junction densities**
β_j	**V**ectors of angles between segments at junction j
H_β	**Entropy of road segment orientation**

Table 1: Summary of the features computed from road networks

2.3 Urban Region Features

We focus on the last four features in Table 2. These features enable us to distinguish between rural classes (Villages and Fields) and urban class (Europe), which otherwise were misclassified due to the lack of extracted network information from the small compact urban regions in the images, shown in Figure 3(a) and Figure 3(c). Let Ω and Ω_R be the area of the image and the area of the extracted regions respectively and $L\psi$ and Γ_R be the network length in $\Psi = \Omega - \Omega_R$ and perimeter of the extracted regions respectively.

We define two descriptors, R_A, the extracted region density and $Cf_A = \Omega_R^{-1}\Gamma_R^2$ the extracted region compactness factor. These two features help us to distinguish the Villages class from the rest of the classes: for example, $R_A \approx 1$ for Urban classes and $R_A \approx 0$ for Mountains and Fields classes.

The number of urban regions in an image, the feature R_v, is used to distinguish between complete Urban, Villages, Fields and Mountains. A complete Urban (USA and Europe) will have $R_v = 1$, whereas, a Villages will have $R_v > 1$, and Fields and Mountains will have $R_v = 0$. Another feature $\Delta_\Omega = \Omega_R / L\psi$, the inverse fractional length density, is also computed to separate the Village class from Urban and Mountains and Fields. For complete Urban classes (USA and Europe), $L\psi = 0$, and for Mountains and Field classes $\Psi = \Omega$. Hence for Mountains and Fields classes, $\Delta_\Omega = 0$, while for complete Urban classes, $\Delta_\Omega = \infty$, and for the Village class $0 < \Delta_\Omega < \infty$.We augment these urban region features with the features computed from the graph representation of the road network as described earlier to improve the classification of the geographical environments which otherwise were misclassified due to the loss of information from small dense urban regions.

Notation	Description
Ω	Area of image
Ω_R	Area of extracted regions
L_Ψ	Network length in $\Psi = \Omega - \Omega_R$
Γ_R	Perimeter of extracted regions
R_A	**Region area density :** Ω_R / Ω
Cf_A	**Region compactness factor** Γ_R^2 / Ω_R
R_v	**Number of regions :** $\# R$
Δ_Ω	**Inverse fractional length density :** Ω_R / L_Ψ

Table 2: Summary of features computed for urban areas.

3. CLASSIFICATION

The 32 features (16 features for each network extraction method) described in section 2.2 were computed for a database of 497 SPOT5, 5m resolution images. To provide ground truth, these images were manually classified into the 7 classes described in section 1 representing various kinds of urban and rural environments. Machine classification was done with a five-fold cross validation on the data set, with 80% of data for training and the remaining 20% for testing in each fold. We performed feature selection using a Fisher Linear Discriminant (FLD) analysis (Duda et al., 2000), followed by a SVM linear kernel classification on the selected feature set. The result of the classification is shown in Table 3. The SVM linear kernel classification on the 30-dimensional feature space selected by the FLD shows a mean error of 24.5% with a standard deviation of 2.92%. As can be clearly seen in the confusion matrix Table 3, the Villages class is confused with the Fields class and also there is a slight confusion between the Urban USA and Urban Europe classes. These confusions arise because, as stated above, the road extraction methods fail to detect the fine and densely structured roads present in some images. Table 4 shows the results of classification of the same set of images with 20 selected feature out of 36 features (32 road network features plus the 4 features computed from the segmented urban areas). As can be seen, there is an improvement in the confusion matrix. The Villages class is less confused with the Fields class than before. The SVM linear kernel classification error is drastically reduced from 24.5%, with only road network features to 12.9%, with the combined feature set with a standard deviation of

3.29%. This is due to the fact, that the loss of information from the urban areas is well captured with the structural features described in section 2.3.

	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7
Villages	**0.55**	0.09	0.22	0.05	0.13	0.00	0.00
Mountains	0.10	**0.81**	0.00	0.00	0.05	0.00	0.02
Fields	0.19	0.05	**0.64**	0.05	0.18	0.00	0.00
USA	0.06	0.00	0.04	**0.82**	0.05	0.00	0.02
Europe	0.09	0.05	0.11	0.07	**0.60**	0.03	0.05
Airports	0.00	0.00	0.00	0.01	0.00	**0.97**	0.00
Common	0.00	0.00	0.00	0.00	0.00	0.00	**0.91**

Table 3: Confusion matrix of a SVM linear kernel classification on 497 images with 7 classes with 30 out of 32 features selected by FLD.

	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7
Villages	**0.83**	0.00	0.15	0.00	0.05	0.02	0.03
Mountains	0.04	**0.83**	0.01	0.00	0.00	0.00	0.00
Fields	0.04	0.08	**0.82**	0.01	0.00	0.00	0.01
USA	0.01	0.00	0.00	**0.92**	0.12	0.02	0.01
Europe	0.08	0.04	0.02	0.07	**0.84**	0.02	0.02
Airports	0.00	0.05	0.00	0.00	0.00	**0.96**	0.00
Common	0.00	0.00	0.00	0.00	0.00	0.00	**0.93**

Table 4: Confusion matrix of a SVM linear kernel classification on 497 images with 7 classes with 20 out of 36 features selected by FLD.

4. INDEXING OF LARGE SPOT5 IMAGES

An image is indexed by a set of keywords representing the content of an image. These keywords are usually limited in numbers and are dependent on application scenarios. Classification is often used as a pre-processing step for indexing. A careful indexing of an image database assists efficient retrieval of image content. The workflow of our indexing method is divided into three steps as follows.

4.1 Step 1: The Database

The image database can be viewed as two sets disjointly partitioned to contain images or segmented images in one set and features extracted from images in another set. We will indicate the image set as S_I, and the feature set as S_F. A pointer is used between S_I and S_F to address a image to its associated feature set. The information extracted in terms of structural features from the large image archive of 497 images, each of size 512x512 pixels, categorized into 7 classes are kept in a data file. The off-line process of this data file creation is done only once, and in case of a new entry, the information extracted from this image is augmented with the existing data file. The pointer is appropriately assigned the address of this new entry. This will be used as the "training" set later in the classification task.

4.2 Step 2: The Feature File

The off-line process for the user given a large image is as follows: the large image of size 5120x5120 pixels is automatically divided into non-overlapping image patches each of size 512x512 pixels. During this process an association pointer is asserted from the image patches, defining its spatial position in the large image. The road network extraction, its graph representation and the urban area segmentation methods are applied in parallel on the image set (100 images). The structural features from the graph representation and the urban areas are stored in a file. The images are *a priori* randomly labelled with classes from 1 to 7. This will be later used as a "testing" test against the above defined "training" set in the classification task.

4.3 Step 3: The Classification

In many satellite image classification works, the *a priori* information about the class label configuration is available and it is very essential and crucial to combine this information into the classification process to obtain a reliable answer. Standard SVM do not provide any estimation of the classification confidence and thus do not allow us to comprehend any *a priori* information. Probabilistic SVM provides us with a solution as to construct a classifier to produce a posterior probability P(class = c|input) which allows us to take a quantitative decision about the classification (Platt, 1999). In this work we used a one-vs-rest Gaussian SVM classifier with σ=10. The choice of the Gaussian standard deviation, σ, which controls the width of the kernel is hard to assert in practical situations. In this study we considered the kernel value which gave us the least training error.

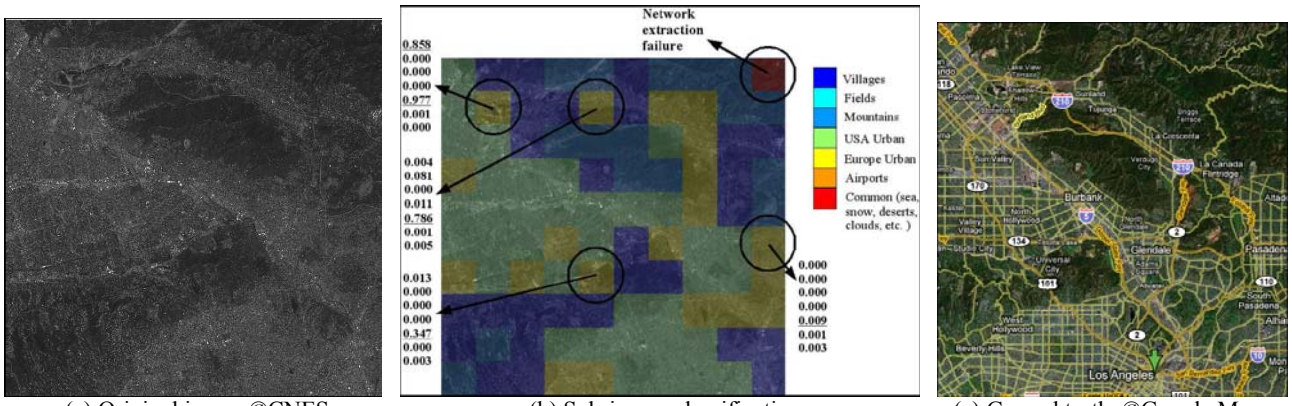
The results of the probabilistic SVM output can be interpreted as follows: the classifier output should be a calibrated posterior probability. First the SVM is trained and then the parameters A and B of an sigmoid function (see Equation 1) are estimated from the training set (f_i, y_i) to map the output of the SVM into probabilities. The predicted label of an image is the one with the largest probability value. The large SPOT5, 5m resolution image, Figure 4(a) of Los Angeles is well classified with a classification accuracy of about 85%.

$$P(y = c|f) = \frac{1}{1 + \exp(A f + B)} \quad (1)$$

The classification image resulting from the probabilistic SVM, Figure 4(b), shows that certain areas are classified as Europe urban. This can be explained from the fact that either the classification probabilities are low or they are comparable with the neighbouring classes. The other reason for this is the fact that the network structures in these areas are similar to the one found in many European urban structures. The superimposed image in Figure 4(c) validates the classified regions with ground truths from Figure 4(d).

5. CONCLUSION

Classification of large satellite images with patches of images extracted from them is a novel idea in the sense that the patches considered contain significant coverage of a particular type of geographical environment. Probabilistic SVM provides us with a quantitative analysis of the classification. This method provides a basis for more complex analysis of large satellite images. The effect of overlapping patches on classification is not reported. This may be an interesting study, as it can help to better classify the images. Moreover, image patches of different sizes can also be used to improve the classification performance. Our indexing method with the above mentioned perspectives can be adapted with existing and future image information mining systems for EO archives.



(a) Original image ©CNES (b) Sub-image classification (c) Ground truth ©Google Maps

Figure 4: Large image indexing

ACKNOWLEDGEMENTS

This work was partially funded by the French Space Agency (CNES), ACI QuerySat, the STIC INRIA-Tunisia programme, and EU NoE Muscle (FP6-507752). The data was kindly provided by CNES and by Sup'com, Tunis. The first author would like to thank INRIA for the PhD fellowship.

REFERENCES

Bhattacharya, A., Jermyn, I. H., Descombes, X. and Zerubia, J., 2006. Computing statistics from a graph representation of road networks in satellite images for indexing and retrieval. In: Proc. CompIMAGE - Computational Modelling of Objects Represented in Images: Fundamentals, Methods and Applications, Coimbra, Portugal.

Bhattacharya, A., Roux, M., Maitre, H., Jermyn, I. H., Descombes, X. and Zerubia, J., 2007. Indexing satellite images with features computed from man-made structures on the earth's surface. In: Proc. International Workshop on Content-Based Multimedia Indexing, pp. 244–250.

Cormen, T., Leiserson, C., Rivest, R. and Stein, C., 2001. Introduction to Algorithms. MIT Press and McGraw-Hill.

Daschiel, H. and Datcu,M., 2005. Image information mining system evaluation using information-theoretic measures. EURASIP Journal on Applied Signal Processing 2005(14), pp. 2153–2163.

Desolneux, A., Maison, L. and Morel, J.-M., 2000. Meaningfulalignments. International Journal of Computer Vision 40(1), pp. 7–23.

Dimitrov, P., Phillips, C. and Siddiqi, K., 2000. Robust and efficient skeletal graphs. In: Proc. IEEE Computer Vision and Pattern Recognition (CVPR), Hilton Head Island, USA, pp. 1417–1423.

Duda, R., Hart, P. and Stork, D., 2000. Pattern Classification, 2ed. Wiley-Interscience.

Fischler, M., Tenenbaum, J. and Wolf, H., 1981. Detection of roads and linear structures in low-resolution aerial imagery using a multisource knowledge integration technique. Computer Graphics and Image Processing 15(3), pp. 201–223.

Khatir, L., 1989. Recherche d'algorithmes de localisation de routes dans les images satellite `a haute r´esolution. Application `a l'imagerie SPOT. PhD thesis, Universit´e Paul Sabatier de Toulouse, France.

Lacoste, C., Descombes, X. and Zerubia, J., 2005. Point processes for unsupervised line network extraction in remote sensing. IEEE Trans. Pattern Analysis and Machine Intelligence 27(10), pp. 1568–1579.

Lowe, D., 2004. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision 60(2), pp. 91–110.

Maître, H., 2007. Indexing and retrieval in large satellite image databases. In: Proc. 5th Int. Symp. Multispectral Image Processing and Pattern Recognition (MIPPR).

Platt, J., 1999. Probabilities for support vector machines. In: A. Smola, P. Bartlett, B. Schlkopf and D. Schuurmans (eds), Advances in Large Margin Classifiers, MIT Press, pp. 61–74.

Rochery, M., Jermyn, I. H. and Zerubia, J., 2003. Higher order active contours and their application to the detection of line networks in satellite imagery. In: Proc. IEEEWorkshop Variational, Geometric and Level Set Methods in Computer Vision, at ICCV, Nice, France.

Roux, M., 1992. Recalage d'images multi-sources. Application au recalage d'une image SPOT et d'une carte. PhD thesis, Ecole National Supérieure des Télécommunications, Paris, France.

Schroeder, M., Rehrauer, H., Seidel, K. and Datcu, M., 1998. Spatial information retrieval from remote sensing images: Part iII Gibbs Markov random fields. IEEE Trans. Geoscience and Remote Sensing 36, pp. 1446–1455.

Serendero, M., 1989. Extraction d'informations symboliques en imagerie SPOT: r´eseaux de communication et agglomérations. PhD thesis, Université de Nice, France.

Siddiqi, K., Bouix, S., Tannenbaum, A. and Zucker, S. W., 2002. Hamilton-jacobi skeleton. International Journal of Computer Vision 48(3), pp. 215–231.